

HOSTED BY



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

[www.elsevier.com/locate/ssci](http://www.elsevier.com/locate/ssci)

# Measurement properties of a screening questionnaire of obstructive sleep apnea risk: Little information, great prediction?☆

Paulo Sargento<sup>a</sup>, Victoria Perea<sup>b</sup>, Valentina Ladera<sup>b</sup>, Paulo Lopes<sup>a</sup>, Jorge Oliveira<sup>a</sup>

<sup>a</sup>Universidade Lusófona de Humanidades e Tecnologias and COPELABS, Portugal

<sup>b</sup>Universidad de Salamanca, Spain

## ARTICLE INFO

### Article history:

Received 18 April 2014

Received in revised form

30 May 2014

Accepted 3 June 2014

Available online 18 September 2014

### Keywords:

Sleep Apnea Risk

Questionnaires

Screening

Classical measurement theory

Rasch rating scale model

## ABSTRACT

**Introduction:** Previous research had shown the suitability of several questionnaires predicting the obstructive sleep apnea syndrome. Measurement properties of an online screening questionnaire were studied.

**Methods:** The sample consisted of 184 Portuguese adults (89 men and 95 women); 46 of them were polysomnographically diagnosed with the untreated obstructive sleep apnea syndrome. The participants were assessed with an online questionnaire of sleep apnea risk, from University of Maryland.

**Results:** A principal component factor analysis was performed, revealing a single factor (49.24% of the total variance). Internal consistency was minimally adequate ( $\alpha=0.74$ ). The mean of inter-item correlation was of 0.35 ( $0.12 < r < 0.61$ ), whereas the item-total correlations were considered good ( $0.52 < r < 0.81$ ). The total score for patients was significantly higher than for healthy participants ( $p < 0.000$ ), but no significant statistical differences between severity groups of patients were found ( $p > 0.05$ ).

Furthermore, the ability of the measure in discriminating between healthy subjects and OSA subjects was good. Overall data from the Rasch analysis was consistent with the guidelines of Linacre, scores show good model fit and psychometric adequacy.

**Conclusions:** The measure showed an adequate structural, internal and criterion validity, suggesting this as a useful and effective screening for sleep apnea risk in Portuguese adults.

© 2014 Published by Elsevier B.V. on behalf of Brazilian Association of Sleep.

☆Institution: The study was developed in Universidade Lusófona de Humanidades e Tecnologias and Universidad de Salamanca, and the Obstructive Sleep Apnea diagnosed subjects were assessed in Hospital Santa Marta (Lisbon), Hospital Amadora/Sintra (Amadora) and Hospital da Beira Interior (Covilhã), in Portugal.

Previous presentation: Some of the data were presented in “LXIV Reunión Anual Sociedad Española de Neurología”, 2012/11/23, Barcelona; Poster Session: “Trastornos de la Vigilia y el Sueño”; Presented by Paulo Sargento.

Peer review under responsibility of Brazilian Association of Sleep.

## 1. Background and objective

The Obstructive Sleep Apnea Syndrome (OSAS) [1] is a breathing-related sleep disorder whose prevalence rates appear to be consistent in different populations. A review [2] of research studies from the United States and Europe has indicated a prevalence of 1–5 Caucasian adults (25–28 kg/m<sup>2</sup> of body mass) have an apnea–hypopnea index (AHI) over 5 (mild severity), and 1–15 Caucasian adults have an AHI over 15 (moderate severity). In line with these findings, it is suggested that adults in Western countries have a 5% probability of suffering from OSAS. Comparing the incidence of OSAS among Western and Eastern countries [3], the rates vary from 3.5% (Australia) to 7.5% (India) for men, whereas for women these values ranged from 1.2% (US) to 4.5% (India). However, 93% of women and 82% of men with moderate to severe OSA are not properly diagnosed [4]. Furthermore, the available data on OSAS allows the identification of the following risk profile: male gender (2–3/1), obesity and aging (30–60 years old, increasing in third age), which are more likely to increase the incidence rates in women after menopause [1,5–11].

The diagnosis of OSAS is one of the key problems in this field. Although polysomnography is considered as the gold standard for the diagnosis of OSAS [1], this method is not yet widely available due, probably, to its high cost. Thus, some attempts have been made to develop self-report measures based on symptoms of sleep apnea (e.g., excessive daytime sleepiness, periods of stopping breathing during sleep observed by others), anthropometry (e.g., body mass index, overweight, neck circumference), demographics (gender and age) and the presence of major comorbidities (e.g., blood pressure, stroke), to screening OSAS's risk. Among the most used screening questionnaires are the Epworth Sleepiness Scale [12], the Berlin Questionnaire [13], the STOP [14], the STOP-Bang [15] and the ASA checklist [16].

A systematic review [17] suggests that STOP and STOP-Bang questionnaires for screening of OSAS in the surgical population are more suitable due to their higher methodological quality and easy-to-use features. On the other hand, a meta-analysis [18] suggested differences in the screening of OSAS between non-risk and surgical populations, in which the Berlin Questionnaire and the Sleep Disorders Questionnaire [19] were considered as effective means to discriminate these differences regarding

OSAS, whereas morphometry and combined clinical cephalometry were the most accurate clinical models.

The Sleep Disorders Center of the University of Maryland has developed a screening questionnaire to assess the risk of OSAS, which is freely available online [20]. This instrument consists of five sections (described in the Section [Method and materials](#)), each of them describing the main features of OSAS, including symptoms observed by the patients or by others, anthropometric characteristics, daytime sleepiness, and major comorbidities. Taking into account the clinical relevance of OSAS and the need to have effective screening measures for this disorder, our main objective was to study the measurement properties of a recent scale to assess OSAS, the Questionnaire of Sleep Apnea Risk (QSAR).

The statistical procedures used to study the effectiveness of QSAR in screening OSAS were based on the Classical Measurement Theory (CMT) and Rasch Measurement Theory (RMT).

## 2. Methods and materials

### 2.1. Design and procedures

This study is based on a one-shot design. All subjects were volunteers (not paid) that gave their informed consent to the study objectives. The study was approved by the scientific and ethical committee of the clinical institutions where the subjects were diagnosed and treated for OSAS.

### 2.2. Participants

The sample consisted of 184 Portuguese adults (89 men and 95 women); 46 of them (35 men and 11 women) were polysomnographically (type 1) diagnosed with untreated OSAS, presenting different types of severity (measured by Apnea–Hypopnea Index [AHI], non-positional). OSAS subjects were recruited from three public Portuguese hospitals, while the remaining participants were healthy volunteers without a clinical diagnosis of OSAS (54 men and 84 women), who were recruited from a sample community (universities and companies) based on convenience method.

**Table 1 – Demographical data and AHI.**

	Healthy subjects n = 138		Subjects with mild to moderate OSAS n = 24		Subjects with severe OSAS n = 22		p
	Mean	SD	Mean	SD	Mean	SD	
Age	41,55	8,57	57,17	11,14	56	8,9	.000*
AHI	— <sup>a</sup>	— <sup>a</sup>	< 30		> 30		—
Gender	♂ n(sr)	♀ n(sr)	♂ n(sr)	♀ n(sr)	♂ n(sr)	♀ n(sr)	
	52(–1,6)	84(1,5)	20(2,5)	4(–2,4)	15(1,3)	7(–1,3)	.000*

OSAS: obstructive sleep apnea syndrome; SD: standard deviation; AHI: apnea/hipopnea index.

<sup>a</sup> No polysomnographic study has performed; ♂: masculine; ♀: feminine; sr: standardised residuals.

**Item 1:** How frequently do you experience or have you been told about snoring loud enough to disturb the sleep of others?  
a)Never  
b)Rarely (less than once a week)  
c)Occasionally (1 - 3 times a week)  
d)Frequently (More than 3 times a week)  
Answer \_\_\_\_\_

**Item 2:** How often have you been told that you have "pauses" in breathing or stop breathing during sleep?  
a)Never  
b)Rarely (less than once a week)  
c)Occasionally (1 - 3 times a week)  
d)Frequently (More than 3 times a week)  
Answer \_\_\_\_\_

**Item 3:** How much are you overweight?  
a)Not at all  
b)Slightly (10 - 20 pounds)  
c)Moderately (20 - 40 pounds)  
d)Severely (More than 40 pounds)  
Answer \_\_\_\_\_

**Item 4:** What is your Epworth Sleepiness Scale Score?  
a)Less than 8  
b) 9-13  
c)14 – 18  
d)19 or greater  
Answer \_\_\_\_\_

**Item 5:** Does your medical history include:  
- High blood pressure  
- Stroke  
- Heart disease  
- More than 3 awakenings per night (on the average)  
- Excessive fatigue  
- Difficulty concentrating or staying awake during the day  
Answer \_\_\_\_\_

**Fig. 1 – Questionnaire of Sleep Apnea Risk (Medical Centre of University of Maryland).**

Prior to enrollment, all the subjects completed a short-form sleeping habits questionnaire, in which healthy participants reported not having any diagnosis or symptoms of sleeping disturbances and overweight. Due to high cost and unavailability of most volunteers, polysomnography was not performed to these participants. They were included only if they scored below 9 in the Epworth Sleepiness Scale [12,21]. None of the subjects assessed (healthy and OSAS subjects) was a shift worker, had clinical history of any neurological or psychiatric disorders, or was doing any type of psychotropic medication.

Table 1 shows the basic demographic characteristics of the sample, as well as the AHI for the clinical sample of OSAS.

The comparisons between healthy subjects and OSAS subjects showed statistically significant differences regarding age. Tukey HSD revealed statistically significant differences between healthy participants and OSAS subjects, but not between the groups of different OSAS severity. As for the gender distribution, standardized residuals showed a difference in the gender distribution, particularly in subjects diagnosed with mild to moderate OSAS.

## 2.3. Materials

### 2.3.1. Clinical history

As stated before, a short-form questionnaire was developed to assess the demographic variables, sleeping habits and if there was a previously diagnosed sleeping disorder.

### 2.4. QSAR [20] (Questionnaire of sleep apnea risk; Medical Centre of University of Maryland)

The QSAR, from University of Maryland, consists in 5 items (including the symptoms observed by the patient and by others, anthropometric characteristics, Epworth Sleepiness Scale score and major comorbidities), as shown in Fig. 1.

The four initial items of the scale are scored from 1(a) to 4(d), whereas item five 'previous medical history' is scored from 1 (none), 2 (1 previous clinical condition), 3 (2–3 previous clinical conditions), and 4 ( $\geq 4$  clinical conditions). The total score is computed through the sum of item responses and represents a measure of sleep apnea risk (ranged 5–20, higher results indicate greater sleep apnea risk).

### Epworth Sleepiness Scale [12]

The Epworth Sleepiness Scale consists of 8 items, rated on a scale of 0–3, in which the total score is computed through the sum of item responses. The total score represents a measure of subjective daytime sleepiness (ranged 0–24, higher results indicate greater propensity to fall asleep).

## 3. Results

The methods of CMT to assess the effectiveness of the QSAR were based on descriptive statistics, principal component factor analysis, inter-item correlations, Cronbach's alpha and an ANOVA for the comparison between different severity OSAS patients's groups and healthy participants. Sensitivity and specificity were also calculated. These methods were performed using the SPSS v.20 for Windows.

The item-response theory through RMT was conducted testing the Linacre guidelines [22]. These analysis were performed using the Winsteps 3.80.1 [23].

### 3.1. Classical measurement theory

#### 3.1.1. Descriptive and distribution analysis

Table 2 shows the descriptive statistics for the 5 items and the total QSAR.

As shown in Table 2, no relevant deviations from Normal distribution were observed for items (i.e., only two items showed slightly negative skewness and one item with slightly lower than normal kurtosis). Descriptive statistics showed that observations covered the scale range for each individual item. The number of missing values was negligible from a statistical standpoint.

**Table 2 – Descriptives for the five items and total QSAR score.**

	Mean	SD	Asymetry	Curtosis	Min/ Max	MV
Item 1	2,33	1,12	0,28	– 1,29	1/4	0
Item 2	1,64	1,10	1,37	0,19	1/4	1
Item 3	2,09	1,02	0,51	– 0,91	1/4	1
Item 4	1,52	0,75	1,27	0,84	1/4	1
Item 5	2,02	1,08	0,68	– 0,85	1/4	0
Total QSAR score	9,58	3,58	0,83	– 0,23	4/19	0

SD: Standard Deviation; Min/Max: Minimum/Maximum; MV: Missing Values.

**Table 3 – Component matrix.**

	Component
Item 1	0.81
Item 2	0.79
Item 3	0.58
Item 4	0.53
Item 5	0.76

Extraction method: principal component analysis.

**Table 4 – Descriptives for QSAR total score.**

	Healthy subjects n = 138		Subjects with mild to moderate OSA n = 24		Subjects with severe OSA n = 22	
	Mean	SD	Mean	SD	Mean	SD
SAR	8,16	2,36	14,25	2,97	13,36	3,54

SD: standard deviation.

#### 3.1.2. Structural validity

To study the structural validity of the scale, a principal component factor analysis was performed on the five items of the QSAR.

The initial solution was minimal satisfactory [KMO=0.709,  $\chi^2(10)=206.479$ ,  $p=0.000$ ]. The communalities ranged from 0.38 (for item 4) to 0.66 (for item 1). From the initial solution, one single factor was extracted with eigenvalue greater than 1 according to the Kaiser and Guttman rule (eigenvalue: 2.462), explaining 49.24% of scale variance.

Table 3 displays the component matrix loadings for each item (> 30) for the one dimension solution.

As shown in Table 3, all items have higher loadings (> 0.50) within a single factor. This one-dimensional solution is suitable to describe our data since it is appropriate to describe the construct and the underlying factor structure.

#### 3.1.3. Internal validity

To study the internal validity of the QSAR, item-total correlations with  $r$  Pearson were performed. The results show moderate to strong positive correlations ( $0.52 < r < 0.81$ ) between each individual item and the total scale (all  $p$ 's=0.000). Item 4 is the one that has the lowest correlation with the total scale.

#### 3.1.4. Criterion validity

The criterion validity of the QSAR was tested with an ANOVA was performed to compare subjects diagnosed with different severity groups OSAS subjects vs. healthy participants in QSAR total score. Table 4 depicts mean scores and standard deviations of the QSAR total score in OSAS subjects and healthy participants.

The comparisons between healthy participants and OSAS subjects groups showed statistically significant differences regarding QSAR global score [ $F(2, 181)=82.169$ ;  $p=0.000$ ]. Tukey test revealed statistically significant differences between healthy participants and OSAS subjects, but not between the groups of different OSAS severities.

#### 3.1.5. Reliability

The internal consistency was estimated using Cronbach's alpha method that was performed to study the reliability of the QSAR in evaluating sleeping disturbances. The Cronbach's alpha (0.74) was acceptable for the version of the QSAR with the original five items, even after the possibility of increasing the alpha level when items were removed. The average inter-item correlation was  $r=0.35$  ( $0.12 < r < 0.61$ ),

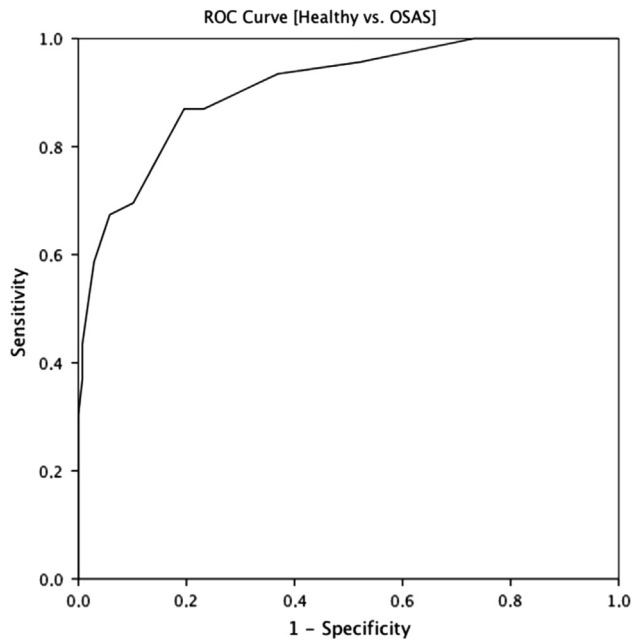


Fig. 2 – ROC curve for Healthy subjects vs OSAS subjects.

Table 5 – Statistics of QSAR.

Category	Observed	B	Infit	Outfit	Step
1	442 (48%)	−1.63	1.14	1.10	–
2	227 (25%)	−0.91	0.94	0.97	−0.73
3	132 (14%)	0.01	0.99	0.96	0.15
4	119 (13%)	0.98	0.91	0.83	0.59
M (SD)					
Item outfit				0.98 (0.24)	
Person outfit				0.98 (0.63)	
Item separation reliability				0.96	
Person separation reliability				0.59	
Cronbach $\alpha$				0.75	
% Items with outfit > 2				0%	
% Persons with outfit > 2				7.61%	
M: mean; SD: standard deviation.					

in which item 3 and item 4 were the most problematic ones according to this analysis.

### 3.1.6. Sensitivity and specificity

The discriminant performance of the QSAR was determined by observing the Receiver Operating Characteristic Curve (ROC). The ROC analysis was performed for the total score of the QSAR. The capability of the scale to distinguish clinical sample from healthy volunteers is shown in Fig. 2.

The Area Under the Curve (AUC) revealed a good discriminant capacity through the total score (AUC=0.91; 95% Confidence Interval ranged between 0.86–0.96). The cut-off score in discriminating among OSAS subjects and healthy subjects was also estimated under the assumption of maximizing the sum of sensitivity and specificity. The best cut-off point for the scale was of 10.5, as shown in Fig. 2.

### 3.2. Rasch measurement theory

Results, through RMT, are consistent with the guidelines of Linacre [22] and all lines of QSAR were overcome successfully. The score statistics is presented in Table 5.

Model fit is adequate: no item outfit is over 2 (severe misfit), the percentage of people with outfit over 2 is small (7.61%) and average outfit values, for items and people, are close to 1 (perfect fit). Furthermore, the score reliability through Item Separation Reliability value (0.96), Cronbach's alpha (0.75), was considered high and Person Separation Reliability value (0.59), indicator that may need more items in the instrument to distinguish between high and low performers in the QSAR.

The graphic representation, in Fig. 3, shows the good functioning of response categories of QSAR, each category has a real probability of being selected by the sample.

## 4. Discussion

In order to increase variability and to prevent floor effects, the statistical analysis was performed for the total sample, including both healthy participants and OSAS subjects. With exception of two items that present a slightly asymmetric distribution, the overall scale distribution is acceptable and did not show any relevant deviations from Normal distribution. From a structural validity point-of-view it was possible to extract an interpretable one-dimensional solution 'sleep apnea risk', which concur with the factor structure of most questionnaires assessing risk of apnea [12–16].

Moreover, this scale reveals an adequate internal validity, which is confirmed by moderate to strong correlations between each of the items and their overall score.

As regards to the criterion validity, the healthy participants differed from the OSAS subjects, but no significant differences were found between severity levels in OSAS groups. Interestingly, the subjects diagnosed with mild to moderate OSAS (AHI<30) showed higher scores on the QSAR than others with severe OSAS (AHI>30). This is unexpected because the AHI, which is a polysomnographic measure of OSAS severity, should be strongly and positively associated with other putative measures (e.g., clinical measures, symptomatic or self-report questionnaires). One possible explanation can be related to the use of new criteria from the AASM [1] for scoring hypopneas that impact on the AHI [24]. The subjects diagnosed with OSAS that comprise our sample were recruited in three different hospitals that use distinct coding systems for hypopneas. On the other hand, it is also known that mild to moderate AHI (11–30) are more variable than extreme AHI (very mild or severe), but also that the use of a single measure of this index in a single polysomnography session is inaccurate and may bias the classification of OSAS severity [25]. Despite we have used the non-positional AHI, which is considered as the more reliable indicator (i.e., with lower variance), it is also known that the percentage of subjects with high variability in AHI assessed exclusively in supine position (position favoring respiratory events) [1] is much higher than that of subjects with high variability in non-positional AHI (which includes all sleeping positions),



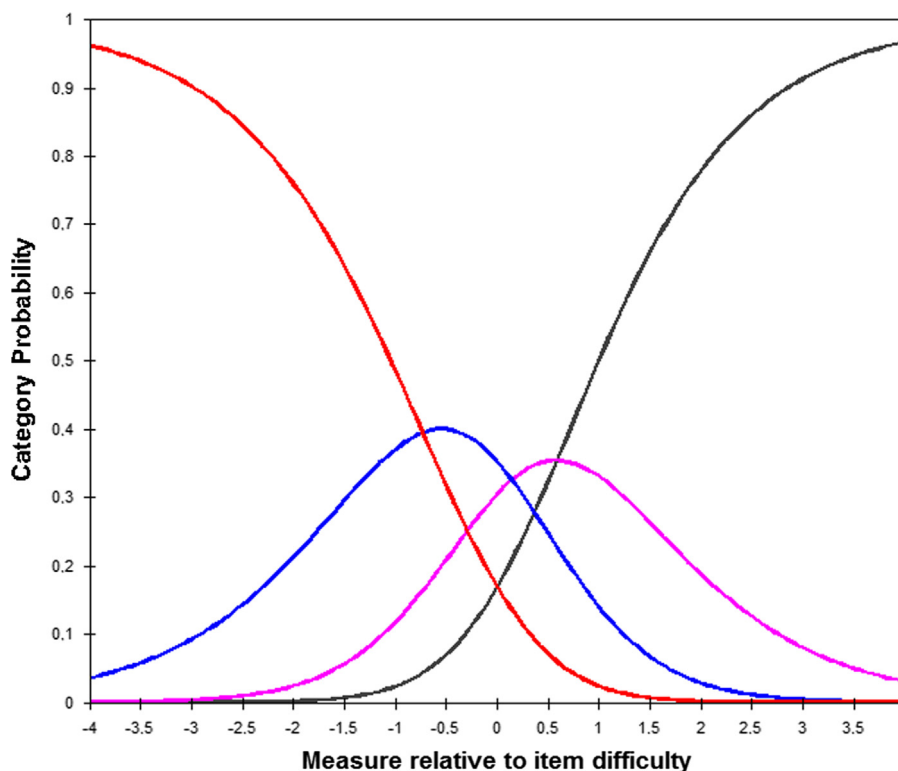


Fig. 3 – Item difficulty.

with exception for cases that the supine position exceeds 35% of night sleep [25]. However, we did not obtain data that allow assessing the amount of supine position per subject.

Regarding the internal consistency of the total scale, estimated through the Cronbach's alpha, the obtained results suggest minimal adequacy of the scale according to Nunally's criteria [26].

We also attempted to study the discriminant capacity of the QSAR in discriminating OSAS subjects from healthy individuals. To accomplish this goal, a ROC analysis was conducted on the total score of the scale. The data indicated a good discriminant performance of the scale. The cut-off score was also estimated under the same statistical procedure. The best cut-off point with adequate level of sensitivity and specificity in discriminating subjects with OSAS is of 10.5. The probability of discrimination a true positive (OSAS; i.e., sensitivity) was, in average, 87%, whereas the probability of discrimination a true negative (without OSAS; i.e., specificity) was, in average, 80%. More particularly, a total score in the QSAR of 10.5 or more may be indicative of OSAS. The obtained data for sensitivity and specificity are encouraging when compared with the most commonly used screening methods for OSAS [16].

The RMT suggests that the standard version of the QSAR is consistent with the theory guidelines [22], scores show good model fit and psychometric adequacy. The main indicators reveal an adequate adjustment between persons and measure.

In sum, both the CMT and RMT suggest suitability of the QSAR to the Portuguese population. The five items that comprise main indicators of OSAS (snoring, stopping breathing, overweight, daytime sleepiness and usual comorbid conditions) can generate relevant information for predicting

OSAS in a fast and simple way. Thus, the QSAR (which includes the Epworth Sleepiness Scale score) provides a useful and effective tool in the first line diagnosis of OSAS in Portuguese adults.

The main limitations to our conclusions were due to the polysomnographic assessments. One of these limitations can be related to differences in the AHI scoring system due to evaluations that were obtained from different polysomnographic centers, which may have an impact on criterion validity to a greater extent than on other statistical procedures in which the overall sample was used. Another limitation that was already mentioned is related to the lack of polysomnographic assessments in our healthy sample of volunteers that may have also contributed to false negative rates. Thus, in future studies, an evaluation in the same center of polysomnography is strongly advised to increase coherence in assessment, and, whenever possible, it is also recommended to include polysomnography for OSAS screening.

### Conflicts of interest

None.

### Authorship contribution

Paulo Sargento (design, background, data collection, statistics and writing main paper); Victoria Perea (writing of abstract and background, and revision); Valentina Ladera (writing of abstract and background, and revision); Paulo Lopes (statistics and results writing); Jorge Oliveira (statistics, results writing and

translation). A translation revision of an English native speaker was taken.

## REFERENCES

- [1] AASM. The International Classification of Sleep Disorders. Diagnostic and Coding Manual. (ICSD.) 2nd Revision. Rochester Minn: American Academy of Sleep Medicine in association with European Sleep Research Society. Japanese Society of Sleep Research and Latin American Sleep Society; 2005.
- [2] Caples SM, Gami AS, Somers VK. Obstructive sleep apnea. *Ann Intern Med* 2005;142(3):187–97.
- [3] Punjabi NM. The epidemiology of adult obstructive sleep apnea. *Proc Am Thorac Soc* 2008;5(2):136–43.
- [4] Young R, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of Sleep Apnea Syndrome in middle-aged men and woman. *Sleep* 1997;20(9):705–6.
- [5] AASM. The International Classification of Sleep Disorders. Diagnostic and Coding Manual, (ICSD) Revised, Rochester Minn: American Academy of Sleep Medicine in association with European Sleep Research Society. Japanese Society of Sleep Research and Latin American Sleep Society; 2001.
- [6] Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. *Am J Respir Crit Care Med* 2002;165(9):1217–39.
- [7] Bixler EO, Vgontzas AN, Ten Have T, Tyson K, Kales A. Effects of age on sleep apnea in men, I: prevalence and severity. *Am J Respir Crit Care Med* 1998;157(1):144–8.
- [8] Bixler EO, Vgontzas AN, Ten Have T, Tyson K, Kales A. Prevalence of sleep-disordered breathing in women: effects of gender. *Am J Respir Crit Care Med* 2001;163(3 Pt 1):608–13.
- [9] Daltro CHC, Fontes FHO, Santos-Jesus R, Gregório PB, Araújo LMB. Síndrome da apnéia e hipopnéia obstrutiva do sono: associação com obesidade, gênero e idade. *Arq Bras Endocrinol Metab* 2006;50(1):74–81.
- [10] Durán J, Esnaola S, Rubio R, Iztueta A. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30–70 yr. *Am J Respir Crit Care Med* 2001;163(3 Pt 1):685–9.
- [11] Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328(17):1230–5.
- [12] Johns M. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991;14(6):540–5.
- [13] Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Ann Intern Med* 1999;131(7):485–91.
- [14] Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108(5):812–21.
- [15] Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth* 2012;108(5):768–75.
- [16] Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, et al. Validation of the Berlin questionnaire and American Society of Anesthesiologists checklist as screening tools for obstructive sleep apnea in surgical patients. *Anesthesiology* 2008;108(5):822–30.
- [17] Abrishtami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anesth* 2010;57:423–38.
- [18] Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology* 2009;110(4):928–39.
- [19] Douglass AB, Bornstein R, Nino-Murcia G, Keenan S, Miles L, Zarcone V, et al. The sleep disorders questionnaire. I: Creation and multivariate structure of SDQ. *Sleep* 1994;17(2):160–7.
- [20] University of Maryland Medical Center. Sleep Apnea Risk. (<http://umm.edu/programs/sleep/health/quizzes/sleep-apnea>); [accessed 13.02.14].
- [21] Johns M. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. *J Sleep Res* 2000;9:5–11.
- [22] Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3(1):85–106.
- [23] Linacre, J.M. A user's guide to Winsteps Ministep – Rasch-model computer programs. Winsteps.com, Chicago; 2013.
- [24] Ruehland WR, Rochford PD, O'Donoghue FJ, Pierce RJ, Singh P, Thornton AT. The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. *Sleep* 2009;32(2):150–7.
- [25] Westbrook PR, Levendowski DJ, Zavora T, Scarfeo D, Berka C, Popovic D. Night to night variability of in-home sleep studies—is one night enough? *Sleep* 2007;30(Abstract Suppl);30(Abstract Suppl) (A188).
- [26] Nunnally JC, Bernstein IH. 3th ed. *Psychometric Theory*, 264–265. USA: McGrawHill; 762.